NORMALIZING GENETIC REPORTER ASSAYS: APPROACHES AND CONSIDERATIONS FOR INCREASING CONSISTENCY AND STATISTICAL SIGNIFICANCE

BY TRISTA SCHAGAT, PH.D., AILEEN PAGUIO, M.S., AND KEVIN KOPISH, B.S., PROMEGA CORPORATION

Genetic reporter assays are used to study DNA sequences and cellular processes that control gene expression. In a typical reporter assay, cells are transfected with a vector that contains the sequence of interest cloned upstream of a reporter protein-coding sequence. Reporter activity is used as an indicator of the ability of the test sequence to regulate gene expression under the experimental conditions. Reporter activity is compared between different vector constructs (e.g., deletion analysis of promoter sequences; see reference 1) or treatment conditions (e.g., screening for G-protein-coupled receptor pathway modulators; see reference 2). Because variables such as cell number and transfection efficiency can have an unwanted effect on the magnitude of reporter expression, reporter data should be normalized. This article discusses approaches for normalization and highlights some key considerations for successful data analysis.

What is Normalization?

Normalization is a process by which data are corrected for factors other than those being directly tested in the experiment. To normalize reporter data, the reporter activity in a particular sample is divided by a second value specific to the same sample. The primary purpose of normalization is to remove sample-to-sample variability caused by factors other than those being tested in the experiment. These factors can include variabilities in cell plating and transfection efficiency, pipetting inconsistencies, and toxicity. Data from each sample is normalized prior to making comparisons between test groups. Data normalization reduces variability and allows data comparisons to be made with greater confidence.

What Methods are used for Normalization?

Various methods have been used for normalization including normalization to total protein content, total ATP content or cell number, and normalization to a control reporter vector. The total protein content and control reporter vector methods are compared in Table 1. Protein normalization can tighten reporter assay results and may be useful when using stably transfected cells. However, most reporter assays are performed using transiently transfected cells, and significant variability can be introduced during transfection. For this reason, vector normalization is recommended for transiently transfected cells.

Vector normalization is accomplished by cotransfection of a control vector, often referred to as an internal vector control, along with the test vector. The internal vector control has a constitutively active promoter driving expression of a second reporter protein. Control reporter activity correlates to the amount of DNA transfected into the cells and the general ability of the cells to express protein. Reporter activity from this internal control is assayed along with the test reporter and used to normalize the test reporter data. In this way, test reporter data are normalized for transfection variability from well-to-well.

What are the Advantages of Vector Normalization?

Vector normalization controls for differences in transfection efficiency between samples. Figure 1 shows two samples with vastly different transfection efficiencies (compare data for no normalization). Despite this difference, data normalized by an internal vector control are still comparable.

By factoring in transfection efficiency, vector normalization reduces data variability and can give differences between test groups greater statistical significance. Figure 2 shows the effect of the normalization method on analysis of three different promoter constructs. When comparing samples with large differences in reporter activity, such as constructs A

Table 1. Comparison of Normalization Methods.

	Normalization Method		
	None	**Protein**	**Vector**
Calculation	None	Reporter Activity / Total Protein	Reporter Activity / Control Reporter Activity
Assays	Reporter	Reporter and Protein	Two Reporters
Ideal Use(s)	Estimations Some Stable Transfections	Stably Transfected Cells	Transiently Transfected Cells
Removes Variability Due To	Not Applicable	Cell Number	Transfection Efficiency

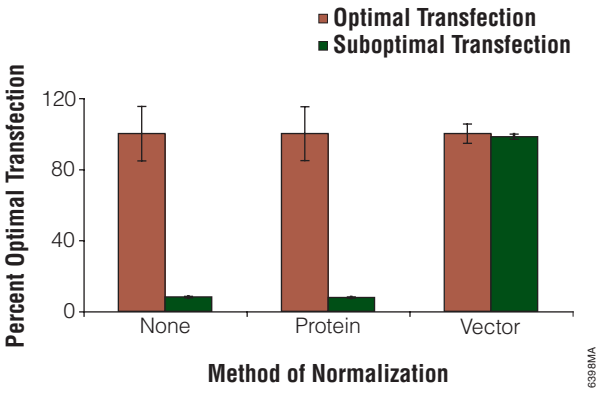

Figure 1. Effect of transfection conditions on reporter results analyzed using different normalization methods. HEK 293 cells were transfected with pGL4.13[*luc2*/SV40] expressing firefly luciferase and pGL4.74[*hRluc*/TK] expressing *Renilla* luciferase. Transfections were performed using both optimal and suboptimal lipid:DNA ratios (indicated on the graph as Optimal and Suboptimal Transfection conditions). Firefly and *Renilla* luciferase activities were measured for each sample using the Dual-Luciferase® Reporter Assay System (Cat.# E1960). Protein concentrations were determined using the Coomassie® Plus Bradford Reagent (Pierce). Firefly luciferase data were either not normalized (None), normalized to total protein (Protein), or normalized to *Renilla* luciferase activity (Vector). Data represent the average ± standard deviation of triplicate samples and are expressed as a percent of the optimal transfection for each normalization condition.

and C, the vector normalization method gives a greater fold change compared to the other analysis methods. The coefficient of variation, which gives a measure of data variability, was significantly lower for vector-normalized results.

Vector normalization also offers greater assay convenience. Luciferases are the most commonly used genetic reporters because luciferase activity assays have broad dynamic range and high sensitivity. This makes them ideal for processing many samples without the need to test multiple sample dilutions or to prepare large quantities of cells. (For a detailed discussion on choosing luciferase reporters and assays, see reference 3.) For vector normalization, activity of two luciferases, firefly and *Renilla*, can be measured in the same cells or lysate aliquot (Table 2). For protein normalization, samples must be split between the luciferase assay (Table 3) and the protein assay. Because the detergents found in lysis buffers can interfere with protein assays, care must be taken to choose an appropriate protein assay (4).

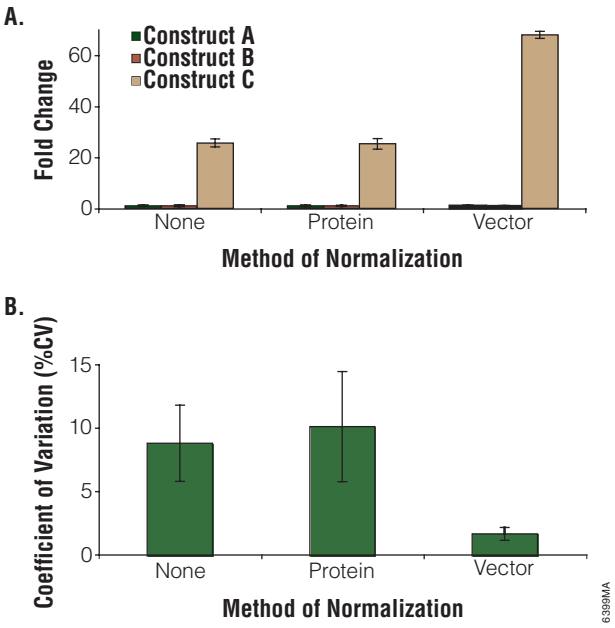

Figure 2. Effect of normalization method on data significance and variability. pGL4.10[*luc2*] (firefly luciferase) reporter vector containing one of three different promoters (constructs A, B, or C) was cotransfected into HEK 293 cells with the pGL4.74[*hRluc*/TK] (*Renilla* luciferase) control vector. Firefly and *Renilla* luciferase activities and protein concentrations were determined as described in the legend to Figure 1. Firefly luciferase data were either not normalized (None), or normalized to total protein (Protein) or *Renilla* luciferase activity (Vector). **Panel A**. Firefly luciferase activity for each normalization method was calculated as fold change compared to Construct A. Data represent the average ± standard deviation of triplicate samples. **Panel B**. The coefficient of variation (standard deviation/average) was calculated for the normalized construct data, and the average ± standard deviation of all the constructs is shown for each normalization method.

How do I Choose an Internal Control Vector?

The internal vector control must encode a reporter other than that used in the test construct. Typically firefly luciferase is used as the test reporter and *Renilla* luciferase as the control reporter. Although there is no technical reason not to reverse these designations, there are more *Renilla* vectors available with different promoters, making *Renilla* the convenient internal control choice (Table 4).

It is important to choose a promoter for the control vector that is compatible with the experimental conditions. The ideal promoter will give low to medium reporter expression and consistent expression under the experimental conditions being tested. The optimal promoter must be determined empirically. The most commonly used promoters include TK, SV40 and CMV.

Table 2. Dual-Luciferase Reporter Assays.

	Dual-Luciferase® Assay	Dual-Glo™ Assay
Format	NH	H
Sample Process	Bench-scale	Batch
Number of Steps	5	2
Sensitivity	Higher	Lower
Firefly Signal Half-Life	~9 minutes	~2 hours
Renilla Signal Half-Life	~2 minutes	~2 hours
Precision	High	High
Cell Lysis Time	~10 minutes maximum	~15 minutes maximum

NH = Non-Homogeneous (lysate is created first)
H = Homogeneous (just add reagent to cells in culture)

Table 3. Single-Luciferase Reporter Assays.

	Bright-Glo™ Reagent	Steady-Glo® Reagent	Luciferase Assay Reagent
Format	NH or H	NH or H	NH
Sample Process	Continuous	Batch	Bench-scale
Number of Steps	1	1	4
Sensitivity	Highest	Lower	High
Firefly Signal Half-Life	~30 minutes	~5 hours	~12 minutes
Precision	High	High	High
Cell Lysis Time	~2 minutes maximum	~5 minutes maximum	variable

NH = Non-Homogeneous (lysate is created first)
H = Homogeneous (just add reagent to cells in culture)

How do I Optimize Transfections to Include the Internal Vector Control?

The optimal amount of control vector to use in cotransfections is the minimum amount that gives significant reporter activity above background (background is measured in samples transfected with only the test vector). This must be determined empirically. Using optimized transfection conditions for a single reporter (for general guidance see reference 5), cotransfect varying amounts of control vector. Typically a ratio of 10:1 test vector:control vector is used; however, as little as 100:1 ratio of test:control vector may be sufficient. The amount of vector needed depends on the control promoter and the cell line.

Figure 3 gives an example of the effect of internal vector control concentration on expression of both test (firefly) and control (*Renilla*) reporters. It is important to ensure that the

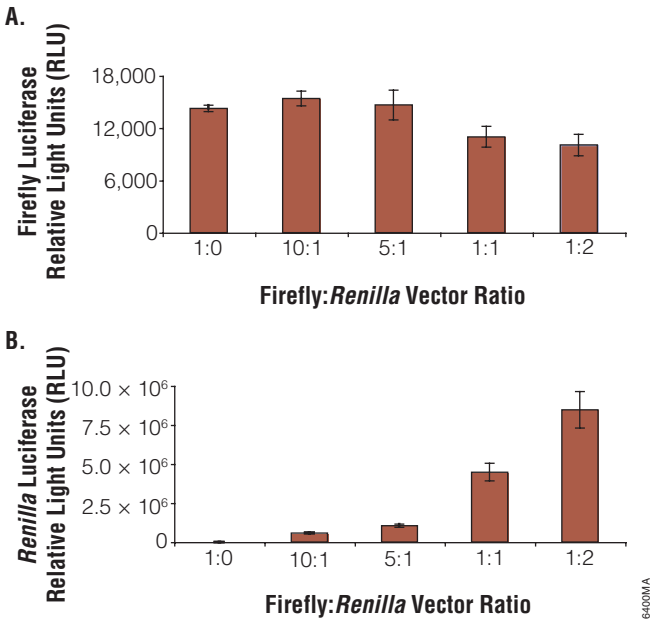

Figure 3. Effect of varying concentrations of internal control (*Renilla* luciferase) vector on test reporter (firefly luciferase) activity. CHO cells were transfected with a constant amount of pGL3 Control Vector expressing firefly luciferase (50ng) and varying amounts (0-100ng) of the internal control vector, phRL-SV40 (expressing *Renilla* luciferase). Total DNA concentration was held constant using carrier DNA. Firefly (**Panel A**) and *Renilla* (**Panel B**) luciferase activities were assayed using the Dual-Luciferase® Reporter Assay System. Data represent the average ± standard deviation of triplicate samples.

presence of the internal control does not affect test reporter activity. As shown by the 1:1 and 1:2 ratios in Figure 3, too much internal control vector will cause a decrease in expression of the test reporter. For this reason, minimal concentrations of the control vector are used.

How are Vector Normalization Calculations Performed?

A single experiment includes identical transfections in triplicate for each test group. Each sample is normalized by dividing the test reporter activity by the control reporter activity. Triplicate samples are then averaged. (Note: Averaging triplicate test and control activities then dividing the averages will give the same result.) This is done for each test group. Table 5 gives an example for an experiment using firefly luciferase as the test reporter and *Renilla* luciferase as the internal control.

Although data are assayed in triplicate, this should not be considered n=3 for drawing scientific conclusions. True replicates are done on separate days with independent samples. The results shown in Table 5 give n=1, and the experiment should be repeated a minimum of 3 times.

Table 4. Common pGL4 Reporter Vectors.

Vector	Multiple Cloning Region	Reporter Gene	Reporter Gene Promoter
pGL4.10[*luc2*][a,b]	Yes	*luc2* (firefly)	No
pGL4.13[*luc2*/SV40][a,b]	No	*luc2* (firefly)	SV40
pGL4.70[*hRluc*][a,c]	Yes	*hRluc* (*Renilla*)	No
pGL4.73[*hRluc*/SV40][a,c]	No	*hRluc* (*Renilla*)	SV40
pGL4.74[*hRluc*/TK][a,c]	No	*hRluc* (*Renilla*)	HSV-TK
pGL4.75[*hRluc*/CMV][a,c,d]	No	*hRluc* (*Renilla*)	CMV

For a complete listing of all pGL4 Vectors, go to: **www.promega.com/pGL4/**

How are Normalized Data Expressed?

The data for reporter assays can be expressed in a number of ways. The following equation can be used to determine the normalized fold change in activity between test groups:

$$\Delta \text{ Fold Activity} = \frac{\text{Average (Firefly/}\textit{Renilla}\text{) from Sample B}}{\text{Average (Firefly/}\textit{Renilla}\text{) from Sample A}}$$

In the example shown in Table 5, each construct is compared to the activity of Construct A. This calculation gives a relative difference in activity between test constructs within a single experiment. The normalized fold changes in activity from each experiment are then averaged together, and the statistical significance determined.

Data may also be expressed as percent activity by multiplying the fold change in activity by 100. Alternatively, the average ratios (Firefly/*Renilla*) can be used for the comparisons between days. Calculating relative differences is not necessary but can be useful.

Regardless of your approach, it is essential that the calculations used be clearly indicated. Minimal manipulation of data is often best, and if you are unsure of the validity of your comparisons, consult with a statistician.

Summary

Data normalization should be considered for all genetic reporter assays. Because of the variability inherent in transient transfections, normalization to an internal vector control should always be used. The ideal control vector will give low to medium expression and consistent results under all test conditions. Normalization is performed for each sample prior to making comparisons between test groups. Using vector normalization, variability is reduced, consistency is increased, and greater significance between samples can be achieved.

Table 5. Sample Data From Transfection Experiments Using the Dual-Luciferase® Reporter Assay System.

Day 1: Test Replicate	Luciferase Activity (RLU)		Ratio (F/R)	Normalized Fold Change in Activity
	Firefly (F)	*Renilla* (R)		
Construct A				
1	27711050	10058796	2.75	
2	25644674	8931007	2.87	
3	29384322	10448575	2.81	
			2.81	**1.00**
Construct B				
1	26580010	11206405	2.37	
2	25141158	10424419	2.41	
3	30456170	12614460	2.41	
			2.40	**0.853**
Construct C				
1	752839872	3877950	194	
2	692411968	3611589	192	
3	669367104	3585187	187	
			192	**67.8**

References

1. Figueiredo, M.S. and Brownlee, G.G. (1995) *J. Biol. Chem.* **270**, 11828–38.
2. Paguio, A. *et al.* (2006) *Cell Notes* **16**, 22–25.
3. Wood, K. (2004) *Cell Notes* **8**, 2–6.
4. Betz, N. (2000) Promega *eNotes* online www.promega.com/enotes/applications/ap0029_print.htm
5. Chapter 12: Bioluminescence Reporters, *Protocols and Applications Guide* (2006) Promega Corporation www.promega.com/paguide/chap12.htm